Expert Systems with Applications 37 (2010) 6093–6101

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Development of an adaptive neuro-fuzzy classifier using linguistic hedges: Part 1

Bayram Cetisli *

Department of Computer Engineering, University of Süleyman Demirel, Isparta 32260, Turkey

ARTICLE INFO

Keywords:
Linguistic hedge (LH)
Adaptive neuro-fuzzy classifier (ANFC)
Scaled conjugate gradient (SCG) algorithm
Computing with words
Pattern recognition

ABSTRACT

In this study, the development of an adaptive neuro-fuzzy classifier (ANFC) is proposed by using linguistic hedges (LHs). The LHs that are constituted by the power of fuzzy sets introduce the importance of the fuzzy sets for fuzzy rules. They can also change the primary meaning of fuzzy membership functions to secondary meaning. To improve the meaning of fuzzy rules and classification accuracy, a layer, which defines the adaptive linguistic hedges, is added into the proposed classifier network. The LHs are trained with other network parameters by scaled conjugate gradient (SCG) training algorithm. The tuned LH values of fuzzy sets improve the flexibility of fuzzy sets, this property of LH can improve the distinguishability rates of overlapped classes. The new classifier is compared with the other classifiers for different classification problems. The empirical results indicate that the recognition rates of the new classifier are better than the other fuzzy-based classification methods with less fuzzy rules.

1. Introduction

Fuzzy systems, which are built on fuzzy rules, have been successfully applied to various classification tasks (Jang, Sun, & Mizutani, 1997; Joshi, Ramakrishman, Houstis, & Rice, 1997; Lin, Yeh, Liang, Chung, & Kumar, 2006; Mitra, De, & Pal, 1997; Nozaki, Ishibuchi, & Tanaka, 1996; Simpson, 1992; Zadeh, 1978). These systems depend on linguistic rules, which are provided by experts or the rules, and are extracted from a given training dataset by a variety of methods like clustering methods (Jang et al., 1997). The fuzzy systems can be constituted with neural networks, and such systems are called as neuro-fuzzy systems (Jang, 1993; Juang & Lin, 1998; Kasabov, 2001; Kasabov & Song, 2002; Nauck & Kruse, 1999). The neuro-fuzzy classifiers define the class distributions and show the input–output relations (Chatterjee & Siarry, 2007; Marin-Blazquez & Shen, 2002; Nauck, 2003; Sun & Jang, 1993), whereas the fuzzy systems describe the systems using natural language. Neural networks are employed for tuning or training the system parameters in neuro-fuzzy applications.

In the early 1970s, Zadeh (1972) introduced a class of powering modifiers, which defined the concept of linguistic variables and hedges. He proposed the computing with words as an extension of fuzzy sets and logic theory (Zadeh, 1996, 1999). The linguistic hedges (LHs) change the meaning of primary terms value. Many researchers have contributed to the computing with words and to the LH concepts with their theoretical studies (De Cock & Kerre, 2004; Huynh, Ho, & Nakamori, 2002; Rubin, 1999; Türkşen, 2004). However, these contributions have been rarely used to solve real-world problems and applications. Chatterjee and Siarry (2007) employed the LH concepts to propose a particle swarm optimization aided neuro-fuzzy classifier (LHBNFC). In the LHBNFC, an LH-based fuzzification layer is added to the network and the LHs are used to modify the piecewise membership subfunctions (Chatterjee & Siarry, 2007). Also, Casillas et al. proposed a fuzzy-genetic inference system based on Mamdani model by applying LHs and power parameters to membership functions (Casillas, Cordon, Del Jesus, & Herrera, 2005). In their study, LHs and power parameters are defined for membership functions, separately. The concept of extended hedge algebras and their application in approximate reasoning were discussed by Ho and Wechler (1992). Modifying the existing LH models, Novak (1996) proposed a horizon-shifting model of LHs, by which the membership function (MF) can be shifted and its steepness modified. Huang, Chen, and Liu (1999) designed a current mode circuit LH for adaptive fuzzy logic controllers, while Zadeh (1975) studied on the approximate reasoning by using the LHs. Banks (1994) used the hedge operations to better qualify and emphasize the crisp variables to mix crisp and fuzzy logic in applications. Bouchon-Meunier (1992) investigated several interesting properties of LHs such as, being compatible with simple symbolic rules, avoiding computations and being compatible with the fuzzy logic, enhancing the comparison of various available fuzzy implications and managing gradual rules in the context of deductive rules. Liu, Chen, and Tsao (2001) also designed an adaptive fuzzy logic controller based on LH concepts.

In literature, a lot of fuzzy classifiers are proposed to solve classification problem (Chakraborty & Pal, 2004; Chatterjee & Siarry, 2007; Kasabov & Song, 2002; Lin et al., 2006; Marin-Blazquez &

* Tel.: +90 246 211 1392; fax: +90 246 237 0859.
E-mail address: bcetisli@mmf.sdu.edu.tr

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.02.108

Shen, 2002; Mitra et al., 1997; Nauck, 2003; Nozaki et al., 1996; Rutkowski & Cpalka, 2003; Shilton & Lai, 2007; Simpson, 1992; Sun & Jang, 1993). However, some of the problems have not been solved. One of them is to discriminate overlapped classes with high reliability. Overlapping classes can significantly decrease the classifier performance. Generally, to overcome this problem, the input space is projected into a new space (Jang et al., 1997). In these cases, the meaning of the original features could be lost. Nozaki and Ishibuchi used the rule weights to solve this problem (Nozaki et al., 1996). Instead of these solutions, some features of input space that cause overlapping can be weakened in classification. When the fuzzy classification rules are investigated, it can be seen that some fuzzy sets affect the classification success as negative due to their overlapping. Sometimes, in fuzzy-based classifications changing the width of the MFs is impossible due to other proximate MFs. The above-mentioned problems can be slightly overcame within an adaptive neuro-fuzzy classifier (ANFC) using the LH concept. And also, most of the fuzzy-based classifiers used are classical fuzzy if–then rules. The other aim of this study is to improve the meaning of the classical fuzzy rules. Therefore, in this study, the development of an ANFC is proposed using LHs. A layer is added to the neuro-fuzzy classifier to indicate the effect of LHs. The empirical studies show that LHs can improve or keep the classification success of the neuro-fuzzy classifier. Especially, the proposed neuro-fuzzy classifier simplifies the distinguishability of overlapping classes. Discriminative features should be used for classification, instead of using all the features, as some of them can cause overlapping among the classes. The selection of discriminative features is given in Part II.

This paper is organized as follows: in Section 2, the descriptions of linguistic variables and hedges are given. The structure of adaptive neuro-fuzzy classifier with linguistic hedges (ANFC-LHs) and its network operations are defined in Section 3. Experimental studies and conclusion are given in Sections 4 and 5, respectively.

2. Linguistic variables and LHs

As pointed by Zadeh (1975), linguistic variables and terms that are more close to human thinking, which bring up importance more than certainty, are seen in everyday life. For that reason, words or linguistic terms are used for modeling the human thinking systems (Liu et al., 2001; Zadeh, 1971). Before using the LHs in the ANFC, the definition of linguistic variable and LHs, and the mathematical representation of LHs in fuzzy logic should be given.

2.1. Linguistic hedges

LHs are special linguistic terms by which other linguistic terms are modified. "Very," "more or less," "fairly," and "extremely" are given as examples of LHs (Jang et al., 1997). For example, A^s = "very young" secondary linguistic term can be produced from the primary linguistic term A = "young" by using LHs (Banks, 1994; Jang et al., 1997; Türkşen, 2004).

2.2. Representation of LHs

An LH or modifier is any operation that changes the meaning of any linguistic term (Banks, 1994; Jang et al., 1997). Let A be a continuous linguistic term for input variable x with MF $\mu_A(x)$. Then A^s is interpreted as a modified version of the original linguistic term expressed as

$$A^s := \{(x, (\mu_A(x))^p) | x \in X\}, \tag{1}$$

where p denotes the linguistic hedge value of the linguistic term A. Two major modifier operations are commonly used in scientific literatures. One of them is concentration (Jang et al., 1997):

$$\mathrm{CON}(A) := A^2. \tag{2}$$

The other operation is dilation (Jang et al., 1997):

$$\mathrm{DIL}(A) := A^{0.5}. \tag{3}$$

Conventionally, CON(A) and DIL(A) are the results of applying hedges "very" and "more or less" to the linguistic term A, respectively. However, there are different and constant LH definitions in literature, such as "very very" ($p = 4$), "quite" ($p = 1.25$), "a little less" ($p = 0.75$) (Banks, 1994; Chatterjee & Siarry, 2007; Jang et al., 1997; Türkşen, 2004). Modified linguistic terms of A for different p values {0, 0.5, 1, 2} and for two type membership functions are shown in Fig. 1. If the LH value is smaller than zero ($p < 0$), the direction of the MF is to change, and obtained membership grades are

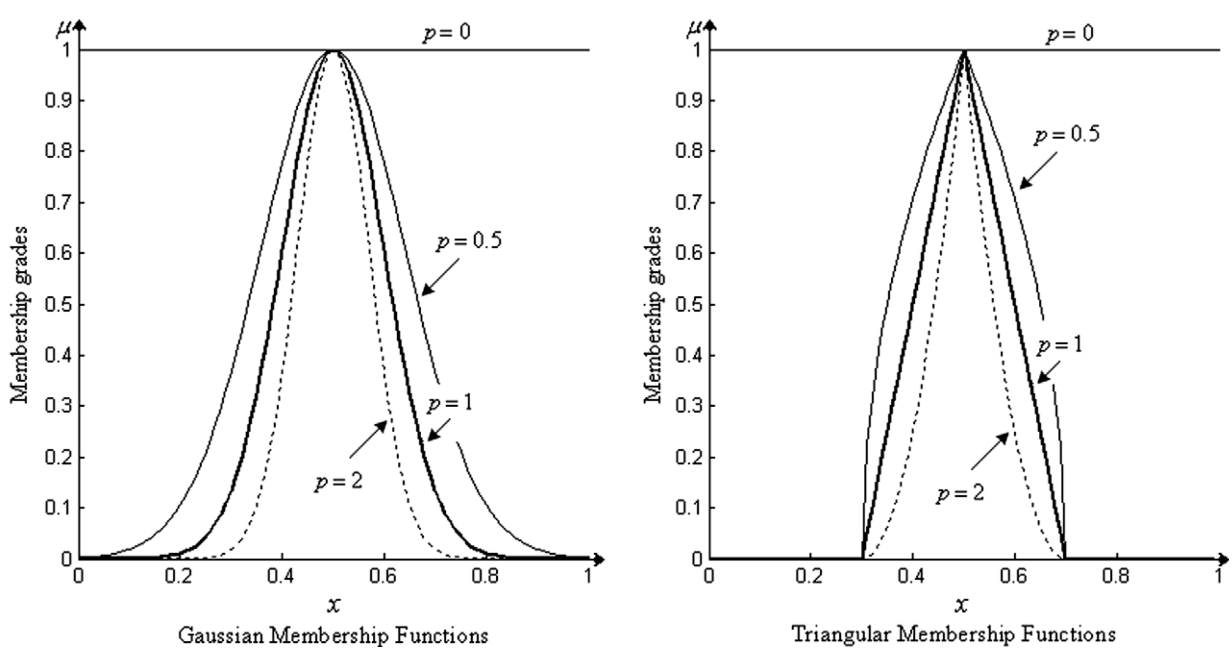

Fig. 1. Interpretation of modified linguistic terms of A that are obtained by different p values.